United States Patent [19]

Berkley et al.

[11] **3,784,747**

[45] **Jan. 8, 1974**

[54] **SPEECH SUPPRESSION BY PREDICTIVE FILTERING**

[75] Inventors: **David Arthur Berkley**, New York, N.Y.; **Olga Mary Mracek Mitchell**, Summit, N.J.

[73] Assignee: **Bell Telephone Laboratories, Incorporated**, Murray Hill, N.J.

[22] Filed: **Dec. 3, 1971**

[21] Appl. No.: **204,509**

[52] **U.S. Cl.** ..... **179/1 HF**, 179/1 P, 179/1 FS

[51] **Int. Cl.** **H04m 1/20**, H04b 15/00

[58] **Field of Search** ..... 179/1 P, 1 VC, 1 HF, 179/1 F, 1 FS, 1 SA, 15.55 R, 100.2 K, 81 B, 170.8; 324/473–476

[56] **References Cited**

UNITED STATES PATENTS

3,177,489	4/1965	Saltzberg	325/476
3,631,520	12/1971	Atal	179/1 SA
3,133,990	5/1964	Seeley	179/1 P
3,644,674	2/1972	Mitchell	179/1 P
3,601,549	8/1971	Mitchell	179/1 FS
3,603,744	9/1971	Krasin	179/81 B

Primary Examiner—William C. Cooper
Assistant Examiner—Jon Bradford Leaheey
Attorney—C. E. Graves

[57] **ABSTRACT**

Speech signal energy from an undesired source is suppressed by extracting from the undesired signal a delay parameter and a gain parameter. These parameters control a delay and gain network through which both the desired and undesired signals are routed. The delayed, amplified signal, when applied to a subtractor along with the undelayed signal, suppresses the undesired signal. The invention is applied to several hands-free telephony situations and to the suppression of one of two speakers in a room.

11 Claims, 10 Drawing Figures

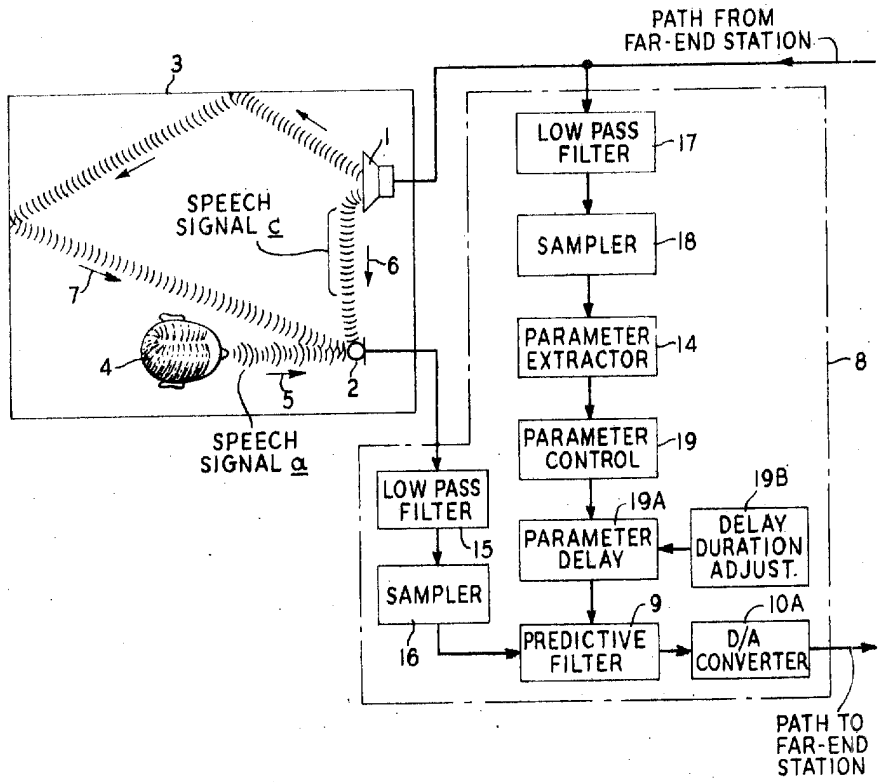


FIG. 1

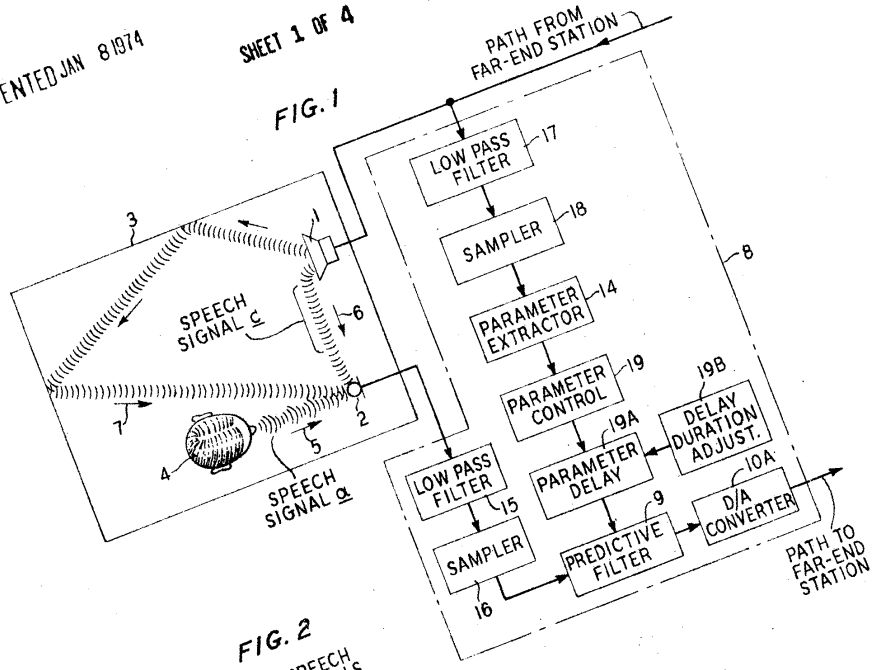


FIG. 2

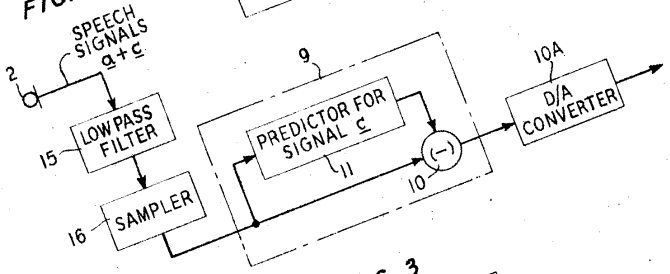


FIG. 3

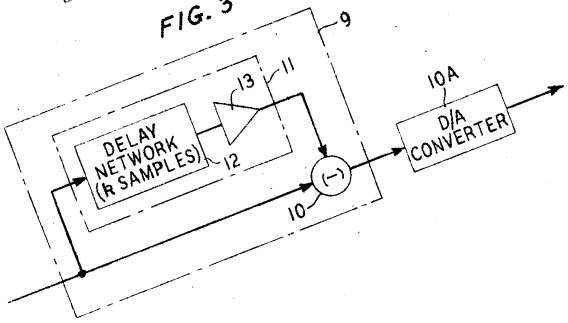


FIG. 4

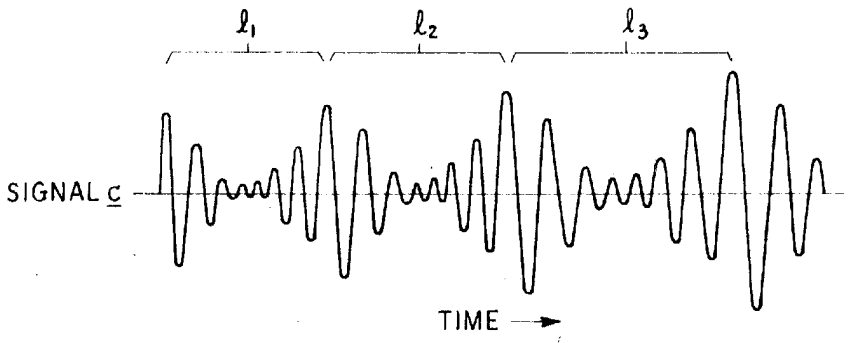


FIG. 5

SPECTRAL MAGNITUDE VS. FREQUENCY
FOR PREDICTIVE FILTER OF FIG. 3

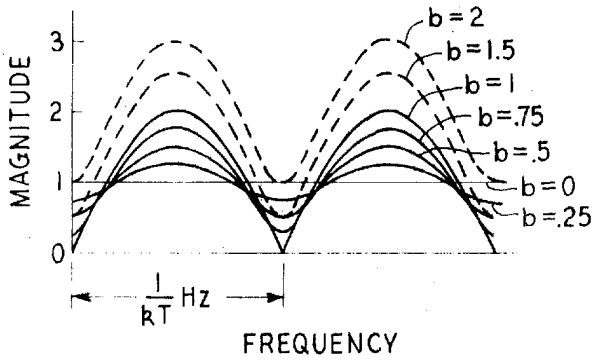
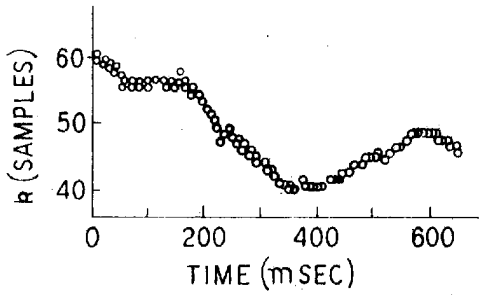


FIG. 6

DELAY PARAMETER k
FOR A TYPICAL VOICED SEGMENT



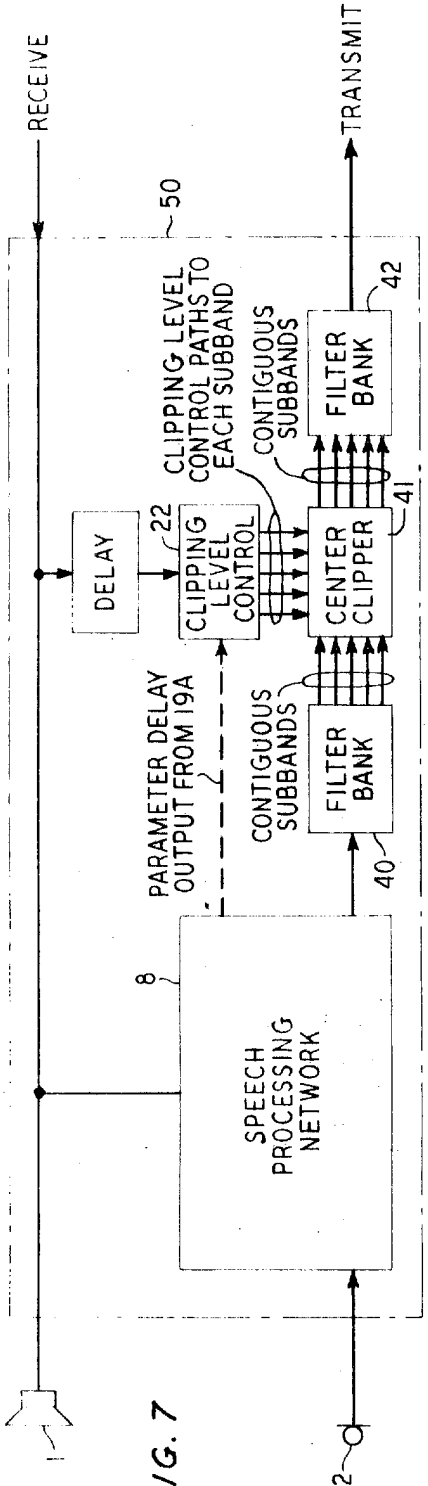


FIG. 7

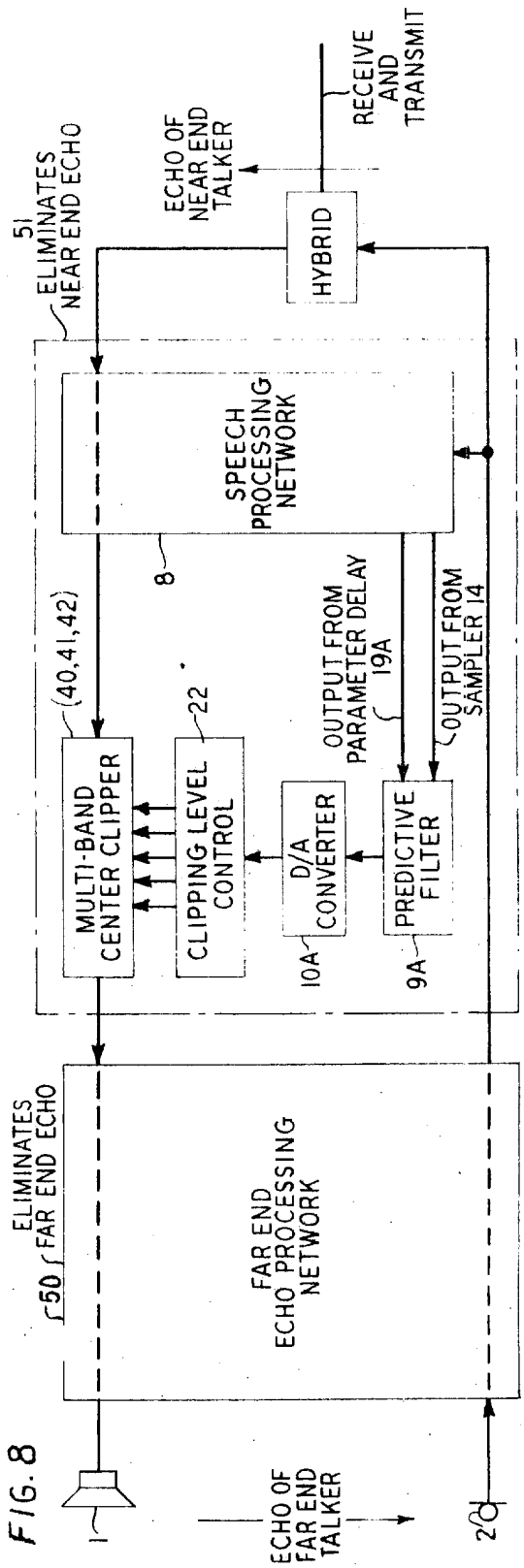


FIG. 8

FIG. 9

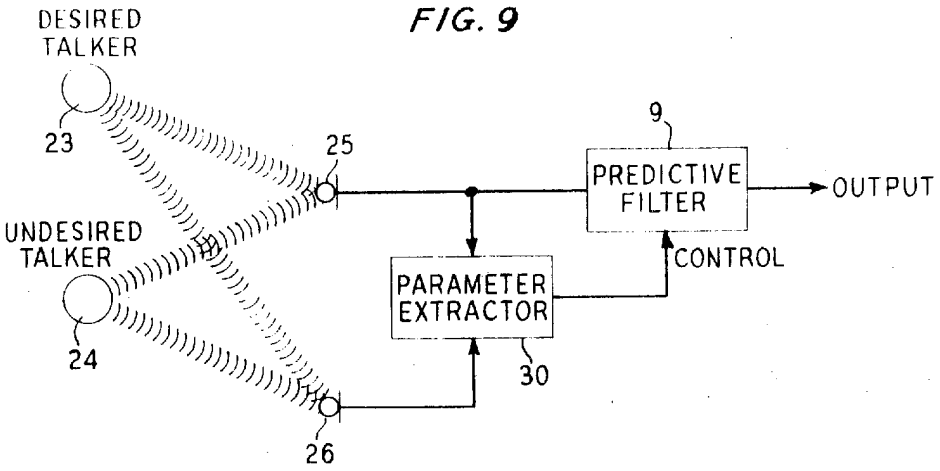
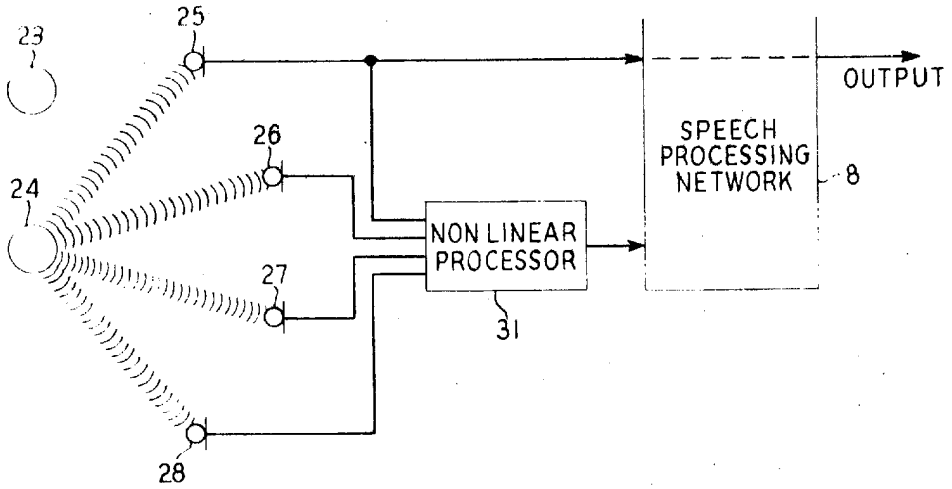


FIG. 10



SPEECH SUPPRESSION BY PREDICTIVE FILTERING

FIELD OF THE INVENTION

This invention relates to speech signal processing, and in particular to reducing the energy content of that part of a composite speech signal attributable to an undesired source.

BACKGROUND OF THE INVENTION

In telephony and elsewhere, it often happens that speech from a source the listener wishes to hear is seriously impaired in intelligibility by speech from a second, undesired source. Numerous expedients to reduce the effects of the second source have been proposed. These involve relative enhancement of the desired speech signal, rendering the undesired signal relatively unintelligible or reducing the energy of the undesired signal. Regardless of the approach, the result typically has been that the desired signal is more intelligible than it would be in the absence of the processing. The "hands-free" telephone well exemplifies this problem of conflicting speech sources, because its electroacoustic speaker constitutes a potential source of undesired signal at the microphone of the same station.

Accordingly, a general object of the invention is to reduce the energy from an undesired speech source in a composite signal containing desired speech.

Another object of the invention is to suppress an undesired speech signal in an electronic communications channel.

A specific object of the invention is to render a desired speech signal relatively more intelligible despite the presence of undesired speech energy.

A particular inventive object is to achieve the foregoing objects in the hands-free telephony situation.

Another specific object of the invention is to avoid voice switching functions and thus enable full-time duplex operation of a hands-free telephone channel.

Yet another inventive object is to distinguish one talker from another nearby talker and to suppress speech signals from one of them.

SUMMARY OF THE INVENTION

The invention is grounded in the general recognition that an unwanted speech signal can be rejected on the basis of its speech parameters.

A discussion of certain speech parameters is found in the patent application of B. S. Atal, Ser. No. 753,408, filed Aug. 19, 1968 now Pat. No. 3,631,520 and assigned to applicants' assignee. A "predictive coding" technique for reducing transmission bandwidth needs is therein disclosed by Atal in which an estimate of the present value of a speech sample is made based on a known corresponding past value. From these data, a difference or error signal is generated, and transmitted to a remote receiving station along with certain predictor parameters. At the remote receiving station the entire signal is reconstituted from the error signal, using the predictor parameters.

It has been realized that the generic process as represented by the Atal disclosure can be rearranged so as to substantially eliminate a given undesired voice signal.

The basic concept contemplated by the present invention is to extract, from the undesired signal, a gain parameter and a delay parameter. These parameters

control a delay and gain network through which both desired speech signals and the undesired signal are routed. The delay is approximately equal to the current duration of the pitch period of the undesired speech. The gain is calculated, in accordance with one of several possible formulas, so as to bring the delayed unwanted signal to the amplitude level of the present value of the unwanted signal. Alternatively, in a technically less complex embodiment, the gain is set equal to 1. In either case, the network output is then subtractively applied to the unwanted signal or to any composite signal containing the unwanted signal. The process may be carried out in analog or digital fashion.

Advantageously however, the process is carried out by sampling techniques where the signal is sampled at a rate of, for example, 6 kHz that results in 30 to 60 samples per pitch period. The number of samples in a pitch period will vary in accordance with the pitch frequency.

In one embodiment pursuant to the invention, speech from the loudspeaker of a hands-free telephone set impinging either directly or reverberatively on the set's microphone, can be largely removed from the microphone output. The reverberant signal as well as the direct signal is suppressed because the unwanted speech parameters do not vary rapidly during voicing.

In another embodiment pursuant to the invention, speech from, for example, two talkers in the same room is detected by a multiplicity of microphones, and the speech of one talker is suppressed using speech parameters determined by combining the outputs of the microphones.

It will be apparent that the rearrangement of the Atal process constitutes in one aspect a filter; and more specifically, a comb filter with minima at the pitch frequency (and harmonics hereof) of the undesired speech. This distinguishes the predictive filter of the present invention from a conventional echo canceler which merely replicates a reverberant signal and subtractively applies the replica to the composite signal.

The invention and its further objects, features, and advantages will be readily discerned in detail from a reading of the description to follow of illustrative embodiments.

BRIEF DESCRIPTION OF THE DRAWING

FIG. 1 is a communications network schematic block diagram containing a hands-free telephone and an inventive embodiment;

FIG. 2 is a schematic block diagram of the inventive predictive filter;

FIG. 3 is a schematic block diagram further delineating the inventive predictor;

FIGS. 4–6 are graphs depicting various characteristics of the predictor;

FIGS. 7 and 8 are two further embodiments of the invention in a communications network containing hands-free telephones; and

FIGS. 9 and 10 are schematic diagrams of the invention as applied to suppression of speech from talkers in a room.

DETAILED DESCRIPTION OF INVENTIVE EMBODIMENTS

Hands-free Telephone Situations

In the first inventive embodiment, a hands-free telephone loudspeaker 1 and microphone 2 present in a re-

verberative enclosure **3** are shown in FIG. **1** connected to the speech processor of the present invention. Usually, the desired speech signal input to microphone **2** is from source **4**, the near-end talker, whose signal denoted *a* travels mainly the direct path **5** and also reverberative paths not shown. Loudspeaker **1** which broadcasts the far-end talker signal, is a source of undesired input to microphone **2** either via the direct path denoted **6** or reverberative paths illustrated by path **7**. The far-end talker direct path speech signal is denoted *c*.

The speech processing network **8** in FIG. **1** consists of what will be called a predictive filter **9** connected in the microphone **2** output circuit. As seen in FIGS. **2** and **3**, filter **9** consists of two parallel legs. The first leg is a predictor **11** which may be a network consisting of a delay network **12** and an amplifier **13**. The second leg is a direct shunt path. Both legs are connected to a subtractor **10**. The predictor **11** is controlled in a manner to be described, by a parameter extractor **14** connected in the loudspeaker **1** circuit.

Pursuant to one embodiment, the invention is carried out digitally. A low pass filter **15** advantageously 3 kHz and a 6 kHz sampler **16** are serially connected in the output circuit of microphone **2**. Similarly, a low pass filter **17** and a sampler **18** are in shunt relation to the loudspeaker **1** input circuit and serially connected to parameter extractor **14**.

A waveform representing the far-end talker signal *c* is illustrated in FIG. **4**. Because a speech signal is redundant—i.e., the signal changes little in shape and length of pitch period from one pitch period to the next—the present form or value of signal *c* can be estimated by a linear prediction based on a past value of signal *c*.

The signal *c* of FIG. **4** is shown made up of speech in consecutive pitch periods l_1 l_2 l_3, etc. Inherently, the speech signals in adjacent pitch periods of signal *c* are of unequal amplitude. Thus, a gain denoted *b* can be calculated (in a manner to be described) that when applied to the sampled signal of the pitch period l_1 will cause the latter to approximate the sampled signal in the next pitch period l_2. If then the amplified signal of period l_1 is subtractively combined with the signal value of period l_2, the result is the substantial filtering out of the signal *c*. In like manner, if a composite signal $a + c$ containing signal *c* is amplified during period l_1 and subtractively combined with the composite signal $a + c$ during period l_2, the same result obtains.

Thus, in mathematical terms, W_n (the amplitude of sample *n* reaching subtractor **10** via the direct path in predictive filter **9**) is subtractively combined with W_{n-k} (the amplitude of the delayed sample) where *k* is the number of samples in a pitch period. Advantageously, the time window over which the parameters are evaluated is of the order of the pitch period to ensure that sufficient energy is present. A time window of 30 samples at a sampling rate of 6 kHz will include between one-half and all the samples in a given pitch period.

Since speech is only quasi-stationary during voicing, the gain parameter *b* and delay parameter *k* have to be periodically calculated. This is accomplished in the digital parameter extractor **14** pursuant to the teaching of the aforementioned Atal patent application Ser. No. 753,408.

As taught therein, input speech samples from sampler **18** are stored as "frames" of signals. The store content is then fed to an arithmetic unit which is part of parameter extractor **14**, wherein for 30 samples, computational values of correlation X_j are computed as follows:

$$X_j = \frac{\sum_{n=0}^{N} W_n W_{n-j}}{\left[\sum_{n=0}^{N} W_n^2\right]^{1/2}\left[\sum_{n=0}^{N} W_{n-j}^2\right]^{1/2}} \qquad (1)$$

where N can advantageously be in the range 30–60 samples.

The computed values of X_j are then inspected in a peak locating network also part of extractor **14**, to determine the largest value of X_j. The value of *j* is found such that X_j is the maximum of all values of X. This particular value of *j* is the delay parameter, *k*, which is supplied to predictive filter **9** as one parameter. It is seen that *k* is a variable delay and that the maximum value of X_j is X_k. The delay parameter *k* for a typical voiced segment is shown in FIG. **6**.

The gain parameter *b* is calculated by computing circuitry also in parameter extractor **14**, that solves:

$$b = \frac{\sum_{n=0}^{N} W_n W_{n-k}}{\sum_{n=0}^{N} W_{n-k}^2} \qquad (2)$$

The gain parameter *b* likewise is supplied to predictive filter **9**.

The described calculation of delay parameter *k* and gain parameter *b* is but one of several systems by which, from an analysis of the speech energy content in adjacent or substantially adjacent signal segments, parameters may be calculated that when applied to a past signal segment will render the latter closely similar to the shape of the present signal segment.

Thus, incoming speech to loudspeaker **1** is continuously analyzed to extract therefrom an optimum delay parameter, and a gain factor. These parameters are periodically updated as for example, every 5 ms. When no incoming signal to loudspeaker **1** is present, the delay and gain are zero. With incoming signal, the calculated present signal value output of predictor **11** is subtracted from the undelayed, unamplified signal sample representing signals $a + c$.

Reconversion to analog form of the signal in the microphone **2** output circuit is achieved in D/A converter **10A**.

The filter depicted in FIG. **3** and described above has a transfer function in Z transform notation.

$$H(Z) = 1 - bZ^{-k} \qquad (3)$$

The magnitude of the frequency response of a typical embodiment of filter **9** is shown in FIG. **5** using predictor **11** where T is the sampling period.

The frequency response for gain parameter $b \leq 1$ are shown by the solid curves and for gain parameter $b \geq 1$ by the broken line. Since speech is dynamic during voicing, the parameters b and k have to be optimized as stated above, and readjusted periodically as, for example, every 5 ms.

Filtering of the input speech by the calculated parameters results in suppression during voicing of up to 30 dB during voiced segments of the undesired signal c, and an average suppression of about 14 dB of the undesired signal c.

The parameters b and k calculated do not vary smoothly with time. The optimum delay occasionally doubles during voiced segments. Also, during unvoiced segments, the optimum delay varies rapidly over a wide range while the correlation remains relatively low. However, the gains calculated are not negligible during these unvoiced portions. Desired speech a, uncorrelated with the undesired signal c which is to be rejected, is degraded when passed through a filter with these rapidly varying filter parameters, while under such conditions no additional suppression of the unwanted source is accomplished.

To avoid this difficulty, logic is introduced pursuant to one facet of the invention, to prevent undesirable variation of the filter parameters b and k. It was determined that not much suppression was obtained when the correlation X_k was less than 0.85. Consequently gain b is set equal to zero for $X_k < 0.85$. This is achieved by parameter control circuit **19** (FIG. 1) which sets b equal to zero for $X_k < 0.85$. This choice of X_k is a compromise between one as great as possible and one low enough so that all of the voiced segments of speech are suppressed. The resulting suppression during voicing is unchanged while degradation of a second speech is reduced. FIG. **6** shows the variation of delay parameter k during a typical voiced segment.

Since the parameters b and k vary relatively slowly during voiced segments, the predictive filter **9** will be effective in removing part of the reverberant signal as well as the direct sound. Specifically, that part of the reverberant signal that has parameters not greatly different from the filter parameters will be reduced in amplitude.

The foregoing discussion of the invention as applied to hands-free telephony has assumed no separation between loudspeaker **1** and microphone **2.** In practice, however, a significant transit time for the signal c to travel path **6** to microphone **2** is required. It is therefore necesssary to compensate in speech processing network **8** for the loudspeaker-microphone transit time. This is achieved by parameter delay circuit **19A** which is serially connected between the output of parameter control **19** and predictive filter **9.** Parameter delay circuit **19A** advantageously is provided with a delay duration adjustment circuit **19B** with which the delay duration may be set to correspond to the transit time which characterizes each given hands-free telephone.

A combination of a predictive filter with a center-clipping echo suppressor of the type taught in D. A Berkley O. M. M. Mitchell-J. R. Pierce U.S. Pat. No. 3,699,271 which is hereby incorporated by reference, is shown in FIGS. **7** and **8.** This combination is a possible replacement for voice switching presently used for echo and feedback suppression.

FIG. **7** shows a network denoted **50** for eliminating the echo of the far-end talker in a 4-wire hands-free telephone. Like numerals denote items which correspond to counterparts in FIGS. **1–3.** The far-end echo picked up by the microphone **2** from loudspeaker **1** is first reduced in amplitude during voiced segments by a speech processor **8** in the manner described previously. Gain and delay parameters b and k of the far-end speech are measured on the received loudspeaker signal, and the far-end echo component of the microphone signal is reduced by filtering. The remaining far-end signal at the output of the speech processor **8** is then removed by the center-clipping echo suppressor.

As taught in D. A. Berkley et al. U.S. Pat. No. 3,699,271, the received signal is used to set the clipping levels by means of clipping control **22** so as just to remove the echo. The output of D/A converter **10A** is fed to filter bank **40** which comprises plural contiguous band filters in the voice frequency range. In center clipper **41** the signal in each subband from filter **41** is center clipped at a level determined by clipping control **22** which measures in effect the energy level in the received signal within each of the subbands. The output of clipper **41** is filtered in bank **42** which is similar to bank **40.**

In this embodiment, the clipping control **22** is advantageously controlled also by the parameter extractor **14.** Since the echo is reduced by the predictive filter **9** during voicing, the clipping levels can be reduced by substantially the same amount during voicing. Consequently in FIG. **7,** a control signal is shown (dashed line) between the parameter extractor **14** and the clipping control **22,** which causes an attenuation of the input to clipping level control **22** that is equal to the suppression achieved by speech processing network **8.** It will be recognized that optimum performance of clipping level control **22** will be realized by inserting a delay in its input path to compensate for the already mentioned signal transit time between loudspeaker **1** and microphone **2.** With the clipping levels thus reduced during voiced segments, there will be less mutilation of the near-end speech by the center-clipping process.

FIG. **8** shows a circuit for eliminating both the far-end echo (echo of far-end talker caused by acoustic coupling through room acoustics) and near-end echo (echo of near-end talker caused by imperfect hybrid junction) in a 2-wire hands-free telephone. The far-end echo is eliminated by network **50** as described above for FIG. **7.** The near-end echo is eliminated by a similar circuit denoted **51** introduced on the receive side of the local 4-wire network as shown.

An alternative method of adjusting the clipping level control by the parameter extractor **14** via the parameter delay **19A** is shown in circuit **51.** A second predictive filter designated **9***a* is used in circuit **51** to attenuate the clipping level control signal during voiced segments. Thus the clipping levels follow the signal at the input to the narrow band center clipper, i.e., at the output of the predictive filter **9***a*.

Suppression of One of Two Room Speakers

A further embodiment allows the suppression of the speech signal from one of two talkers in a room. FIG. **9** shows the desired speech source **23** and an undesired source **24** both of whose speech signals form the input to microphones **25** and **26.** The undesired source **24** is positioned so that the time delays for direct sound transmission to microphones **25** and **26** are equal. In the output of microphone **25** is predictive filter **9** as

shown in FIG. 2 and previously described. The predictor 11 shown in FIG. 3 is controlled by a two-microphone parameter extractor 30.

Two signals, W_n' from microphone 25 and W_n from microphone 26, enter the parameter extractor 30 wherein an arithmetic unit within the extractor calculates the computational values

$$X_j \quad \sum_{n=0}^{N} W_n W_{n-j} \Big/ \left[\left(\sum_n W_n^2 \right)^{1/2} \left(\sum_n W_{n-j}^2 \right)^{1/2} \right] \quad (4)$$

and

$$Y_j = \sum_{n=0}^{N} W_n' W_{n-j} \Big/ \left[\left(\sum_n W_n'^2 \right)^{1/2} \left(\sum W_{n-j}^2 \right)^{1/2} \right] \quad (5)$$

A peak picking network within the extractor then selects the peaks from X_j and Y_j and a comparator finds the largest value peak which occurs in both sequence X_j and Y_j for the same value of j. This value of j is the delay parameter k for the undesired speech supplied to the predictive filter 9.

An alternative method of extracting the parameters is shown in FIG. 10. Two additional microphones 27 and 28 are positioned so that time delays from desired speaker 24 for direct sound transmission to microphones 27 and 28 are equal to the time delays to microphones 25 and 26. The outputs of all microphones 25–28 are processed by a non-linear processor 31 as described in O. M. M. Mitchell-C. A. Ross-R. L. Wallace, Jr. U.S. Pat. No. 3,644,671, which is hereby incorporated by reference. The output of processor 31 contains the undesired signal and an attenuated and disturbed component of the desired signal. (The outputs may alternatively be added to merely attenuate the desired signal.) The output of the non-linear processor 31 enters the speech processing network 8. The output of microphone 25 is processed by speech processing network 8 which filters out the undesired talker 24 in the manner already described. The presence in the output of the nonlinear processor 31 of a small amount of the desired talker does not significantly affect the delay parameter k but will cause a small error in the evaluation of X_k and b.

It is to be understood that the embodiments described herein are merely illustrative of the principles of the invention. Various modifications may be made thereto by persons skilled in the art without departing from the spirit and scope of the invention.

What is claimed is:

1. Speech processing apparatus for suppressing voiced segments of an undesired speech signal while leaving a desired speech signal intelligible, comprising:

means for deriving an electronic waveform representing the undesired speech signal;

means for deriving an electronic waveform representing a composite signal containing a reverberant version of the undesired speech signal and the desired signal;

means for deriving from the waveform of said undesired speech signal a delay parameter determined from the signal values during an interval embracing a substantial portion of a pitch period of said undesired speech signal;

means for applying said composite speech signal waveform to a summer over a first path;

means for delaying in a second path said composite speech signal waveform by an amount of said delay parameter; and

means for subtractively applying to said summer the delayed said composite speech waveform.

2. Apparatus pursuant to claim 1, further comprising means responsive to the absence of voiced segments of said undesired speech signal for interrupting said second path.

3. Apparatus in accordance with claim 1 further comprising means for deriving from the waveform of said undesired speech signal a gain parameter specifying the amount by which the amplitudes of corresponding values of said undesired speech signal in a past said interval must be respectively adjusted so as to produce a substantial duplicate of the undesired said speech signal of a present said interval; and which further comprises means controlled by said gain parameter for amplifying said delayed composite speech waveform prior to its being subtractively applied to said summer.

4. A communications network comprising:

a hands-free telephone station including a direct acoustic coupling path between the station loudspeaker and microphone, a second remote telephone station, and transmission means interconnecting said stations;

means for deriving—from the incoming signal waveform to said loudspeaker from said second station—a delay parameter representing the duration of an interval embracing a substantial portion of the present pitch period of speech from the remote station; and a gain parameter specifying the amount by which the waveform in a past said interval must be changed in amplitude to substantially correspond to the undesired speech waveform of the present interval;

means for applying the composite signal—consisting of the desired near-end talker signal and the acoustically coupled far-end talker signal from said loudspeaker—in said microphone output to a summer over a first path;

means disposed in a second path for delaying said composite signal by the amount of said delay parameter and for amplifying the delayed composite signal an amount determined by said gain parameter; and

means for subtractively applying the delayed, amplified composite signal to said summer.

5. A communications network pursuant to claim 4 wherein said deriving means comprises:

a signal sampler connected to the circuit of said loudspeaker and operating at a set sampling rate; and

means for computing values of a term X_j in accordance with the relationship

$$X_j = \frac{\sum_{n=0}^{N} W_n W_{n-j}}{\left[\sum_{n=0}^{N} W_n^2 \right]^{1/2} \left[\sum_{n=0}^{N} W_{n-j}^2 \right]^{1/2}}$$

where W_n is the amplitude of a sample n reaching said sampler, and means for finding that value of j such that X_j is the maximum of all values of X_j, the

found value of j constituting said delay parameter.

6. A communications network pursuant to claim 5 wherein said gain parameter deriving means comprises means for computing the value b in accordance with the relationship

$$b=\frac{\sum_{n=0}^{N} W_n W_{n-k}}{\sum_{n=0}^{N} W_{n-k}^2}$$

where W_n is the amplitude of a sample n reaching said sampler, and k is a delay parameter for a voiced segment.

7. A communications network pursuant to claim 6, further comprising means for rendering said gain parameter equal to zero in the absence of voiced segments of the signal in said loudspeaker path from said remote station.

8. A communications network pursuant to claim 7, further comprising means for setting said gain parameter equal to zero in response to values of X_k, corresponding to the maximum computed values of X_j which are less than a critical predetermined value.

9. A communications network pursuant to claim 8, further comprising means for adjustably delaying arrival of said delay and gain parameters at said second path by an amount that compensates for the transit time delay over said direct acoustic coupling path of speech from said remote station.

10. A communications network pursuant to claim 4, further comprising:

filter bank means connected to the output of said summer and comprising plural contiguous subbands;

means for producing—from said remote station speech signal—control signals representative of the incoming speech energy level in each said subband;

means for center-clipping the output of each said filter bank subband a varying amount in response to the concurrent said energy level value; and

means connecting said control signal producing means and said deriving means responsive to voiced portions of signal from said remote station for reducing all said clipping levels.

11. Speech processing apparatus for suppressing speech from one of two talkers in a room comprising:

first and second microphones located equidistant from the first, desired said talker but at unequal distances from the second, undesired said talker;

means for deriving from the two said microphone outputs a first parameter X_j in accordance with the relationship

$$X_j=\sum_{n=0}^{N} W_n W_{n-j} \Big/ \left[\left(\sum_n W_n^2\right)^{1/2}\left(\sum_n W_{n-j}^2\right)^{1/2}\right]$$

and a second parameter Y_j, respectively, calculated in accordance with the relationship

$$Y_j=\sum_{n=0}^{N} W'_n W_{n-j} \Big/ \left[\left(\sum_n W'^2_n\right)^{1/2}\left(\sum W_{n-j}^2\right)^{1/2}\right]$$

where W_n is the speech signal received by said first microphone and W_n' is the speech signal received by said second microphone;

means for selecting the largest value peak from the composite peak values of said parameters X_j and Y_j for the same value of the term j;

means for applying the desired and the undesired said signals from one of said microphones to a summer directly over a first path and alternately over a second path through a network including delay means; and

means for adjusting said delay means as a function of the value of the term j.

* * * * *